APPLICATION UNDER UNITED STATES PATENT LAWS

Atty. Dkt. No. PM 275901/AFG/Z70363/UST
(M#)

Invention: HUMAN CCR-2 GENE POLYMORPHISMS

Inventor (s): SMITH, John Craig
ANAND, Rakesh
MORTEN, John Edward Norris

Pillsbury Madison & Sutro LLP
Intellectual Property Group
1100 New York Avenue, NW
Ninth Floor
Washington, DC 20005-3918
Attorneys
Telephone: (202) 861-3000

This is a:

- ☐ Provisional Application
- ☐ Regular Utility Application
- ☐ Continuing Application
  - ☒ The contents of the parent are incorporated by reference
- ☒ PCT National Phase Application
- ☐ Design Application
- ☐ Reissue Application
- ☐ Plant Application
- ☐ Substitute Specification
  Sub. Spec Filed _____
  in App. No. _____ / _____
- ☐ Marked up Specification re
  Sub. Spec. filed _____
  In App. No _____ / _____

SPECIFICATION

HUMAN CCR-2 GENE POLYMORPHISMS

This invention relates to polymorphisms in the human CCR-2 gene and corresponding novel allelic polypeptides encoded thereby. The invention also relates to methods and materials for analysing allelic variation in the CCR-2 gene, and to the use of CCR-2 polymorphism in the diagnosis and treatment of CCR-2 ligand mediated diseases such as rheumatoid arthritis and other inflammatory diseases.

MCP-1 acts through the CCR-2 receptor (also known as the MCP-1 receptor). MCP-2 and MCP-3 may also act, at least in part, through the MCP-1 receptor.

MCP-1 is a member of the chemokine family of pro-inflammatory cytokines which mediate leukocyte chemotaxis and activation. MCP-1 is a C-C chemokine which is one of the most potent and selective T-cell and monocyte chemoattractant and activating agents known. MCP-1 has been implicated in the pathophysiology of a large number of inflammatory diseases including rheumatoid arthritis, glomerular nephritides, lung fibrosis, restenosis (International Patent Application WO 94/09128), alveolitis (Jones *et al.*, 1992, J. Immunol., **149**, 2147) and asthma. Other disease areas where MCP-1 is thought to play a part in their pathology are atherosclerosis (e.g. Koch *et al.*, 1992, J. Clin. Invest., 90, 772-779), psoriasis (Deleuran *et al.*, 1996, J. Dermatological Science, **13**,. 228-236), delayed-type hypersensitivity reactions of the skin, inflammatory bowel disease (Grimm *et al.*, 1996, J. Leukocyte Biol., **59**,. 804-812), multiple sclerosis and brain trauma (Berman *et al*, 1996, J. Immunol., **156**,. 3017-3023). An MCP-1 inhibitor may also be useful to treat stroke, reperfusion injury, ischemia, myocardial infarction and transplant rejection.

CCR-2 polypeptide is known to exist in 2 isoforms, CCR-2A and CCR-2B, which are alternatively spliced variants of a single CCR-2 gene. The reader is referred to the following publications: Organization and differential expression of the Human Monocyte Chemoattractant Protein 1 Receptor Gene: Evidence for the role of the carboxy-terminal tail in receptor trafficking, LM Wong *et al* J Biol Chem **272**,1038-1045 (1997), see Figure 1 therein in particular; and International patent application WO 95/19436, Charo *et al.*

One known polymorphism is Val64Ile, the reader is referred to the following

publications: A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation, LG Kostrikis *et al* Nature Medicine **4**, 350-353 (1998); and The role of CCR5 and CCR2 polymorphisms in HIV-1 transmission and disease progression, NL Michael *et al* Nature Medicine **3**, 1160-1162 (1997).

The CCR-2 gene has been cloned and published as a EMBL Accession number: U 80924 (5471 bp) and all positions herein of polymorphisms in the coding sequence relate to the position therein unless stated otherwise or apparent from the context.

The genomic sequence of CCR-2 is contained in the BAC (Bacterial Artificial Chromosome) clone, 110P12 (Research Genetics), which is published as EMBL Accession number U95626 (143068 bp). All positions herein of polymorphisms in the promoter region relate to the position therein unless stated or otherwise apparent from the context.

The protein structure of the CCR-2 has been published (Molecular structure of chemokine receptors, R Horuk, Trends in Pharmaceutical Sciences **15**, 159-165 (1994)).

One approach is to use knowledge of polymorphisms to help identify patients most suited to therapy with particular pharmaceutical agents (this is often termed "pharmacogenetics") . Pharmacogenetics can also be used in pharmaceutical research to assist the drug selection process. Polymorphisms are used in mapping the human genome and to elucidate the genetic component of diseases. The reader is directed to the following references for background details on pharmacogenetics and other uses of polymorphism detection: Linder *et al.* (1997), Clinical Chemistry, **43**, 254; Marshall (1997), Nature Biotechnology, **15**, 1249; International Patent Application WO 97/40462, Spectra Biomedical; and Schafer et al. (1998), Nature Biotechnology, 16, 33.

Clinical trials have shown that patient response to treatment with pharmaceuticals is often heterogeneous. Thus there is a need for improved approaches to pharmaceutical agent design and therapy.

Point mutations in polypeptides will be referred to as follows: natural amino acid (using 1 or 3 letter nomenclature) , position, new amino acid. For (a hypothetical) example "D25K" or "Asp25Lys" means that at position 25 an aspartic acid (D) has been changed to lysine (K). Multiple mutations in one polypeptide will be shown between square brackets with individual

mutations separated by commas.

The present invention is based on the discovery of two single nucleotide polymorphisms (SNPs) in the coding sequence of the CCR-2 gene and 11 SNPs in the promoter sequence of the CCR-2 gene.

According to one aspect of the present invention there is provided a method for the diagnosis of a single nucleotide polymorphism in CCR-2 in a human, which method comprises determining the sequence of the nucleic acid of the human at one or more of positions 2385 and 2649 in the coding sequence of the CCR-2 gene as defined by the positions in EMBL ACCESSION NO. U 80924, and/or one or more of positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 and 43018 in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U95626;
and determining the status of the human by reference to polymorphism in the CCR-2 gene..

The term human includes both a human having or suspected of having a CCR-2 ligand mediated disease and an asymptomatic human who may be tested for predisposition or susceptibility to such disease. At each position the human may be homozygous for an allele or the human may be a heterozygote.

The term single nucleotide polymorphism includes single nucleotide substitution, nucleotide insertion and nucleotide deletion which in the case of insertion and deletion includes insertion or deletion of one or more nucleotides at a position of a gene.

In one embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 2385 is presence of C and/or T.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 2649 is presence of G and/or A.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 40915 is presence of A and/or T.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 41047 is the presence or

absence of an insertion of ACA.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 41058 is the presence of C and/or A.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 41507 is the presence of C and/or A.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 41768 is the presence of A and/or T.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 42401 is the presence of A and/or G.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 42598 is presence or absence or an insertion of T.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 42673 is the presence of G and/or A.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 42723 is the presence of C and/or A.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 42874 is the presence of A and/or G.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 43018 is the presence of A and/or T.

The method for diagnosis is preferably one in which the sequence is determined by a

method selected from amplification refractory mutation system and restriction fragment length polymorphism.

Allelic variation at position 2385 consists of a single base substitution from C (the published base), preferably to T. Allelic variation at position 2649 consists of a single base substitution from G (the published base), preferably to A. Allelic variation at position 40915 consists of a single base substitution from A (the published base), preferably to T. Allelic variation at position 41047 consists of the absence of insertion (the published case), preferably to the insertion of ACA. Allelic variation at position 41058 consists of a single base substitution from C (the published base), preferably to A. Allelic variation at position 41507 consists of a single base substitution from C (the published base), preferably to A. Allelic variation at position 41768 consists of a single base substitution from A (the published base), preferably to T. Allelic variation at position 42401 consists of a single base substitution from A (the published base), preferably to G. Allelic variation at position 42598 consists of a single base substitution from the absence of insertion (the published case), preferably to the insertion of T. Allelic variation at position 42673 consists of a single base substitution from G (the published base), preferably to A. Allelic variation at position 42723 consists of a single base substitution from C (the published base), preferably to A. Allelic variation at position 42874 consists of a single base substitution from A (the published base), preferably to G. Allelic variation at position 43018 consists of a single base substitution from A (the published base), preferably to T. The status of the individual may be determined by reference to allelic variation at any one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve or thirteen positions.

CCR-2 ligand antagonist drugs have been disclosed in the following publications: US patent 5707815, University of California; International patent application WO 98/06703, Warner Lambert; and Japanese patent application JP 09309877-A, Teijin Limited.

The test sample of nucleic acid is conveniently a sample of blood, bronchoalveolar lavage fluid, sputum, or other body fluid or tissue obtained from an individual. It will be appreciated that the test sample may equally be a nucleic acid sequence corresponding to the sequence in the test sample, that is to say that all or a part of the region in the sample nucleic acid may firstly be amplified using any convenient technique e.g. PCR, before analysis of allelic variation.

It will be apparent to the person skilled in the art that there are a large number of analytical procedures which may be used to detect the presence or absence of variant nucleotides at one or more polymorphic positions of the invention. In general, the detection of allelic variation requires a mutation discrimination technique, optionally an amplification reaction and optionally a signal generation system. Table 1 lists a number of mutation detection techniques, some based on the PCR. These may be used in combination with a number of signal generation systems, a selection of which is listed in Table 2. Further amplification techniques are listed in Table 3. Many current methods for the detection of allelic variation are reviewed by Nollau *et al.*, Clin. Chem. **43**, 1114-1120, 1997; and in standard textbooks, for example "Laboratory Protocols for Mutation Detection", Ed. by U. Landegren, Oxford University Press, 1996 and "PCR", 2nd Edition by Newton & Graham, BIOS Scientific Publishers Limited, 1997.

Abbreviations:

ALEX™	**Amplification refractory mutation system linear extension**
APEX	**Arrayed primer extension**
ARMS™	**Amplification refractory mutation system**
b-DNA	**Branched DNA**
C-C chemokine	**This is a group of chemokines based on arrangement of the first two of four conserved cysteine residues; the other group is called "C-X-C" chemokines (see Schall (1994) in The Chemokine Handbook, 419, Academic Press, editor Thomson).**
CCR	**C-C chemokine receptor**
CCR-2A and CCR-2B	**are isoforms of the CCR-2 polypeptide produced as splice variants from a single CCR-2 gene**
CMC	**Chemical mismatch cleavage**
bp	**base pair**
COPS	**Competitive oligonucleotide priming system**
DGGE	**Denaturing gradient gel electrophoresis**
FRET	**Fluorescence resonance energy transfer**

LCR	**Ligase chain reaction**
MASDA	**Multiple allele specific diagnostic assay**
MCP-1	**Monocyte chemoattractant protein 1**
NASBA	**Nucleic acid sequence based amplification**
OLA	**Oligonucleotide ligation assay**
PCR	**Polymerase chain reaction**
PTT	**Protein truncation test**
RFLP	**Restriction fragment length polymorphism**
SDA	**Strand displacement amplification**
SNP	**Single nucleotide polymorphism**
SSCP	**Single-strand conformation polymorphism analysis**
SSR	**Self sustained replication**
TGGE	**Temperature gradient gel electrophoresis**

Table 1 - Mutation Detection Techniques

General: DNA sequencing, Sequencing by hybridisation

Scanning: PTT*, SSCP, DGGE, TGGE, Cleavase, Heteroduplex analysis, CMC, Enzymatic mismatch cleavage

* Note: not useful for detection of promoter polymorphisms.

Hybridisation Based

Solid phase hybridisation: Dot blots, MASDA, Reverse dot blots, Oligonucleotide arrays (DNA Chips)

Solution phase hybridisation: Taqman™ - US-5210015 & US-5487972 (Hoffmann-La Roche), Molecular Beacons - Tyagi *et al* (1996), Nature Biotechnology, 14, 303; WO 95/13399 (Public Health Inst., New York)

Extension Based: ARMS™, ALEX™ - European Patent No. EP 332435 B1 (Zeneca Limited), COPS - Gibbs *et al* (1989), Nucleic Acids Research, 17, 2347.

Incorporation Based: Mini-sequencing, APEX

Restriction Enzyme Based: RFLP, Restriction site generating PCR

Ligation Based: OLA

Other: Invader assay

Table 2 - Signal Generation or Detection Systems

Fluorescence: FRET, Fluorescence quenching, Fluorescence polarisation - United Kingdom Patent No. 2228998 (Zeneca Limited)

Other: Chemiluminescence, Electrochemiluminescence, Raman, Radioactivity, Colorimetric, Hybridisation protection assay, Mass spectrometry

Table 3 - Further Amplification Methods

SSR, NASBA, LCR, SDA, b-DNA

Preferred mutation detection techniques include ARMS™, ALEX™, COPS, Taqman, Molecular Beacons, RFLP, and restriction site based PCR and FRET techniques.

Particularly preferred methods include ARMS™ and RFLP based methods. ARMS™ is an especially preferred method.

In a further aspect, the diagnostic methods of the invention are used to assess the efficacy of therapeutic compounds in the treatment of CCR-2 ligand mediated diseases.

Assays, for example reporter-based assays, may be devised to detect whether one or more of the above polymorphisms affect transcription levels and/or message stability.

Individuals who carry particular allelic variants of the CCR-2 gene may therefore exhibit differences in their ability to regulate protein biosynthesis under different physiological conditions and will display altered abilities to react to different diseases. In addition, differences in protein regulation arising as a result of allelic variation may have a direct effect on the response of an individual to drug therapy. The diagnostic methods of the invention may be useful both to predict the clinical response to such agents and to determine therapeutic dose.

In a further aspect, the diagnostic methods of the invention, are used to assess the predisposition and/or susceptibility of an individual to diseases mediated by CCR-2 ligands.

This may be particularly relevant in the development of rheumatoid arthritis and cardiovascular disease and the present invention may be used to recognise individuals who are particularly at risk from developing these conditions.

In a further aspect, the diagnostic methods of the invention are used in the development of new drug therapies which selectively target one or more allelic variants of the CCR-2 gene. Identification of a link between a particular allelic variant and predisposition to disease development or response to drug therapy may have a significant impact on the design of new drugs. Drugs may be designed to regulate the biological activity of variants implicated in the disease process whilst minimising effects on other variants.

In a further diagnostic aspect of the invention the presence or absence of variant nucleotides is detected by reference to the loss or gain of, optionally engineered, sites recognised by restriction enzymes. In the accompanying Example 2 we provide details of convenient engineered restriction enzyme sites that are lost or gained as a result of a polymorphism of the invention.

According to another aspect of the present invention there is provided a human CCR2 gene or its complementary strand comprising a polymorphism corresponding with one or more of positions 2385 and 2649 as defined by the positions in EMBL ACCESSION NO. U 80924 and in which there is a T at position 2385 and an A at position 2649 or a fragment thereof of at least 20 bases comprising at least one of the polymorphisms.

Fragments are at least 17 bases, more preferably at least 20 bases, more preferably at least 30 bases.

According to another aspect of the present invention there is provided a human CCR2 gene or its complementary strand comprising a polymorphism corresponding with one or more positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 and 43018 in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U95626 and in which there is a T at position 40915, an insertion of ACA at position 41047, an A at position 41058, an A at position 41507, a T at position 41768, a G at 42401, the insertion of a T at position 42598, an A at position 42673, an A at position 42723, a G at position 42874 ,or a T at position 43018 or a fragment thereof of at least 20 bases comprising

at least one polymorphism.

Fragments are at least 17 bases, more preferably at least 20 bases, more preferably at least 30 bases.

According to another aspect of the invention there is provided a nucleotide sequence comprising a human CCR2 gene or its complementary strand or an antisense sequence thereto comprising a polymorphism at one or more of: positions 2385 and 2649 as defined by the positions in EMBL ACCESSION NO. U 80924 and in which there is a T at position 2385 and an A at position 2649; or positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 and 43018 in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U95626 and in which there is a T at position 40915, an insertion of ACA at position 41047, an A at position 41058, an A at position 41507, a T at position 41768, a G at 42401, the insertion of a T at position 42598, an A at position 42673, an A at position 42723, a G at position 42874 and a T at position 43018; or a fragment thereof of at least 20 bases comprising at least one polymorphism.

Novel sequence disclosed herein, may be used in another embodiment of the invention to regulate expression of the gene in cells by the use of antisense constructs. To enable methods of down-regulating expression of the gene of the present invention in mammalian cells, an example antisense expression construct can be readily constructed for instance using the pREP10 vector (Invitrogen Corporation). Transcripts are expected to inhibit translation of the gene in cells transfected with this type construct. Antisense transcripts are effective for inhibiting translation of the native gene transcript, and capable of inducing the effects (e.g., regulation of tissue physiology) herein described. Oligonucleotides which are complementary to and hybridizable with any portion of novel gene mRNA disclosed herein are contemplated for therapeutic use. U.S. Patent No. 5,639,595, Identification of Novel Drugs and Reagents, issued Jun. 17, 1997, wherein methods of identifying oligonucleotide sequences that display in vivo activity are thoroughly described, is herein incorporated by reference. Expression vectors containing random oligonucleotide sequences derived from previously known polynucleotides are transformed into cells. The cells are then assayed for a phenotype resulting from the desired activity of the oligonucleotide. Once cells with the desired phenotype have been identified, the sequence of the

oligonucleotide having the desired activity can be identified. Identification may be accomplished by recovering the vector or by polymerase chain reaction (PCR) amplification and sequencing the region containing the inserted nucleic acid material.nucleotide molecules can be synthesized for antisense therapy. These antisense molecules may be DNA, stable derivatives of DNA such as phosphorothioates or methylphosphonates, RNA, stable derivatives of RNA such as 2'-O-alkylRNA, or other oligonucleotide mimetics. U.S. Patent No. 5,652,355, Hybrid Oligonucleotide Phosphorothioates, issued July 29, 1997, and U.S. Patent No. 5,652,356, Inverted Chimeric and Hybrid Oligonucleotides, issued July 29, 1997, which describe the synthesis and effect of physiologically-stable antisense molecules, are incorporated by reference. Antisense molecules may be introduced into cells by microinjection, liposome encapsulation or by expression from vectors harboring the antisense sequence.

According to another aspect of the present invention there is provided a computer readable medium comprising at least one novel polynucleotide sequence of the invention stored on the medium. The computer readable medium may be used, for example, in homology searching, mapping, haplotyping, genotyping or pharmacogenetic analysis or any other bioinformatic analysis. The reader is referred to Bioinformatics, A practical guide to the analysis of genes and proteins, Edited by A D Baxevanis & B F F Ouellette, John Wiley & Sons, 1988. Any computer readable medium may be used, for example, compact disk, tape, floppy disk, hard drive or computer chips.

The polynucleotide sequences of the invention, or parts thereof, particularly those relating to and identifying the single nucleotide polymorphisms identified herein represent a valuable information source, for example, to characterise individuals in terms of haplotype and other sub-groupings, such as investigation of susceptibility to treatment with particular drugs. These approaches are most easily facilitated by storing the sequence information in a computer readable medium and then using the information in standard bioinformatics programs or to search sequence databases using state of the art searching tools such as "GCC". Thus, the polynucleotide sequences of the invention are particularly useful as components in databases useful for sequence identity and other search analyses. As used herein, storage of the sequence information in a computer readable medium and use in sequence databases in relation to

'polynucleotide or polynucleotide sequence of the invention' covers any detectable chemical or physical characteristic of a polynucleotide of the invention that may be reduced to, converted into or stored in a tangible medium, such as a computer disk, preferably in a computer readable form. For example, chromatographic scan data or peak data, photographic scan or peak data, mass spectrographic data, sequence gel (or other) data.

The invention provides a computer readable medium having stored thereon one or a more polynucleotide sequences of the invention. For example, a computer readable medium is provided comprising and having stored thereon a member selected from the group consisting of: a polynucleotide comprising the sequence of a polynucleotide of the invention, a polynucleotide consisting of a polynucleotide of the invention, a polynucleotide which comprises part of a polynucleotide of the invention, which part includes at least one of the polymorphisms of the invention, a set of polynucleotide sequences wherein the set includes at least one polynucleotide sequence of the invention, a data set comprising or consisting of a polynucleotide sequence of the invention or a part thereof comprising at least one of the polymorphisms identified herein. A computer based method is also provided for performing sequence identification, said method comprising the steps of providing a polynucleotide sequence comprising a polymorphism of the invention in a computer readable medium; and comparing said polymorphism containing polynucleotide sequence to at least one other polynucleotide or polypeptide sequence to identify identity (homology), i.e. screen for the presence of a polymorphism.

The invention further provides nucleotide primers which can detect the polymorphisms of the invention.

According to another aspect of the present invention there is provided an allele specific primer capable of detecting a CCR-2 gene polymorphism at one or more of positions 2385 and 2649 in the CCR-2 gene as defined by the positions in EMBL ACCESSION NO. U 80924, and/or one or more of positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 and 43018 in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U95626.

An allele specific primer is used, generally together with a constant primer, in an amplification reaction such as a PCR reaction, which provides the discrimination between alleles

through selective amplification of one allele at a particular sequence position e.g. as used for ARMS™ assays. The allele specific primer is preferably 17- 50 nucleotides, more preferably about 17-35 nucleotides, more preferably about 17-30 nucleotides.

An allele specific primer preferably corresponds exactly with the allele to be detected but derivatives thereof are also contemplated wherein about 6-8 of the nucleotides at the 3' terminus correspond with the allele to be detected and wherein up to 10, such as up to 8, 6, 4, 2, or 1 of the remaining nucleotides may be varied without significantly affecting the properties of the primer.

Primers may be manufactured using any convenient method of synthesis. Examples of such methods may be found in standard textbooks, for example "Protocols for Oligonucleotides and Analogues; Synthesis and Properties," Methods in Molecular Biology Series; Volume 20; Ed. Sudhir Agrawal, Humana ISBN: 0-89603-247-7; 1993; 1st Edition. If required the primer(s) may be labelled to facilitate detection.

According to another aspect of the present invention there is provided an allele-specific oligonucleotide probe capable of detecting a CCR-2 gene polymorphism at one or more of positions 2385 and 2649 in the CCR-2 gene as defined by the positions in EMBL ACCESSION NO. U 80924, and/or one or more positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 , 43018 and in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U95626.

The allele-specific oligonucleotide probe is preferably 17- 50 nucleotides, more preferably about 17-35 nucleotides, more preferably about 17-30 nucleotides.

The design of such probes will be apparent to the molecular biologist of ordinary skill. Such probes are of any convenient length such as up to 50 bases, up to 40 bases, more conveniently up to 30 bases in length, such as for example 8-25 or 8-15 bases in length. In general such probes will comprise base sequences entirely complementary to the corresponding wild type or variant locus in the gene. However, if required one or more mismatches may be introduced, provided that the discriminatory power of the oligonucleotide probe is not unduly affected. The probes of the invention may carry one or more labels to facilitate detection.

According to another aspect of the present invention there is provided a diagnostic kit comprising an allele specific oligonucleotide probe of the invention and/or an allele-specific

primer of the invention.

The diagnostic kits may comprise appropriate packaging and instructions for use in the methods of the invention. Such kits may further comprise appropriate buffer(s) and polymerase(s) such as thermostable polymerases, for example taq polymerase.

In another aspect of the invention, the single nucleotide polymorphisms of this invention may be used as genetic markers in linkage studies. This particularly applies to the polymorphisms at 2385, 41047, 41507 and 42723 because of their relatively high frequency (see below). The CCR-2 gene has been mapped to chromosome 3p21.3-p24 (Samson *et al* (1996), Genomics **36**, 522-526). Low frequency polymorphisms may be particularly useful for haplotyping as described below. A haplotype is a set of alleles found at linked polymorphic sites (such as within a gene) on a single (paternal or maternal) chromosome. If recombination within the gene is random, there may be as many as 2^n haplotypes, where 2 is the number of alleles at each SNP and n is the number of SNPs. One approach to identifying mutations or polymorphisms which are correlated with clinical response is to carry out an association study using all the haplotypes that can be identified in the population of interest. The frequency of each haplotype is limited by the frequency of its rarest allele, so that SNPs with low frequency alleles are particularly useful as markers of low frequency haplotypes. As particular mutations or polymorphisms associated with certain clinical features, such as adverse or abnormal events, are likely to be of low frequency within the population, low frequency SNPs may be particularly useful in identifying these mutations (for examples see: Linkage disequilibrium at the cystathionine beta synthase (CBS) locus and the association between genetic variation at the CBS locus and plasma levels of homocysteine. *Ann Hum Genet* (1998) **62**:481-90, De Stefano V, Dekou V, Nicaud V, Chasse JF, London J, Stansbie D, Humphries SE, and Gudnason V; and Variation at the von willebrand factor (vWF) gene locus is associated with plasma vWF:Ag levels: identification of three novel single nucleotide polymorphisms in the vWF gene promoter. *Blood* (1999) **93**:4277-83, Keightley AM, Lam YM, Brady JN, Cameron CL, Lillicrap D).

According to another aspect of the present invention there is provided a method of treating a human in need of treatment with a CCR-2 ligand antagonist drug in which the method comprises:

i) diagnosis of a single nucleotide polymorphism in CCR-2 gene in the human, which diagnosis comprises determining the sequence of the nucleic acid at one or more of positions 2385 and 2649 in the CCR-2 gene as defined by the positions in EMBL ACCESSION NO. U 80924, and/or at one or more positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 and 43018 in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U 95626,

and determining the status of the human by reference to polymorphism in the CCR-2 gene; and

ii) administering an effective amount of a CCR-2 ligand antagonist .

Preferably determination of the status of the human is clinically useful. Examples of clinical usefulness include deciding which antagonist drug or drugs to administer and/or in deciding on the effective amount of the drug or drugs. CCR-2 ligand antagonist drugs have been disclosed in the following publications: US patent 5707815, University of California; International patent application WO 98/06703, Warner Lambert; and Japanese patent application JP 09309877-A, Teijin Limited.

According to another aspect of the present invention there is provided use of a CCR-2 ligand antagonist drug in preparation of a medicament for treating a CCR-2 ligand mediated disease in a human diagnosed as having a single nucleotide polymorphism at one or more of positions 2385 and 2649 in CCR-2 gene as defined by the positions in EMBL ACCESSION NO. U 80924, and/or at one or more positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 and 43018 in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U 95626.

According to another aspect of the present invention there is provided a pharmaceutical pack comprising a CCR-2 ligand antagonist drug and instructions for administration of the drug to humans diagnostically tested for a single nucleotide polymorphism at one or more of positions 2385 and 2649 in CCR-2 gene as defined by the positions in EMBL ACCESSION NO. U 80924 and/or at one or more positions 40915, 41047, 41058, 41507, 41768, 42401, 42598, 42673, 42723, 42874 and 43018 in the promoter sequence of the CCR-2 gene as defined by the positions in the EMBL ACCESSION NO. U95626.

The invention will now be illustrated but not limited by reference to the following

Examples. All temperatures are in degrees Celsius.

In the Examples below, unless otherwise stated, the following methodology and materials have been applied.

AMPLITAQ™, available from Perkin-Elmer Cetus, is used as the source of thermostable DNA polymerase.

General molecular biology procedures can be followed from any of the methods described in "Molecular Cloning - A Laboratory Manual" Second Edition, Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory, 1989).

Electropherograms were obtained in a standard manner: data was collected by ABI377 data collection software and the wave form generated by ABI Prism sequencing analysis (2.1.2).

Example 1

Identification of Polymorphisms

1. Methods

DNA Preparation

DNA was prepared from frozen blood samples collected in EDTA following protocol I (Molecular Cloning: A Laboratory Manual, p392, Sambrook, Fritsch and Maniatis, 2nd Edition, Cold Spring Harbor Press, 1989) with the following modifications. The thawed blood was diluted in an equal volume of standard saline citrate instead of phosphate buffered saline to remove lysed red blood cells. Samples were extracted with phenol, then phenol/chloroform and then chloroform rather than with three phenol extractions. The DNA was dissolved in deionised water.

Template Preparation

Templates were prepared by PCR using the oligonucleotide primers and annealing temperatures set out below. The extension temperature was 72° and denaturation temperature 94°. Generally 50 ng of genomic DNA was used in each reaction and subjected to 35 cycles of PCR and primary fragment was diluted 1/200 before amplification of secondary fragments. PCR was performed in two stages (primary fragment then secondary fragment) to ensure specific

amplification of the desired target sequence.

EMBL ACCESSION NO. U 80924

Primary Fragment	Forward Oligo	Reverse Oligo	Annealing Temp	Time
1478-3534	1478-1497	3515-3534	55°	30s

Secondary Fragments	Forward Oligo	Reverse Oligo	Annealing Temp	Time
1511-1902	1511-1530	1883-1902	55°	30s
1840-2189	1840-1859	2170-2189	55°	30s
2111-2491	2111-2130	2472-2491	55°	30s
2433-2842	2433-2452	2823-2842	55°	30s

EMBL ACCESSION NO: U95626:

Fragment	Forward Oligo	Reverse Oligo	Annealing Temp	Time
40716-41155	40716-40735	41136-41155	60°	60s
41069-41506	41069-41088	41487-41506	58°	60s
41451-41883	41451-41470	41864-41883	58°	60s
41802-42250	41802-41821	42231-42250	58°	60s
42123-42560	42123-42142	42541-42560	58°	60s
42484-42927	42484-42503	42908-42927	58°	60s
42847-43299	42847-42866	43280-43299	58°	60s

For dye-primer sequencing these primers were modified to include M13 forward and reverse primer sequences (ABI protocol P/N 402114, Applied Biosystems) at the 5' end of the forward and reverse oligonucleotides respectively.

Dye Primer Sequencing

Dye-primer sequencing using M13 forward and reverse primers was as described in the ABI protocol P/N 402114 for the ABI Prism™ dye primer cycle sequencing core kit with

"AMPLITAQ™FS" DNA polymerase, modified in that the annealing temperature was 45° and DMSO was added to the cycle sequencing mix to a final concentration of 5 %.

The extension reactions for each base were pooled, ethanol/sodium acetate precipitated, washed and resuspended in formamide loading buffer.

4.25 % Acrylamide gels were run on an automated sequencer (ABI 377, Applied Biosystems).

2. Results

Novel Polymorphisms

EMBL ACCESSION NO. U 80924

Nucleotide 2385	C/T	Asn (260)	AAC/AAT	AAC	30.6 %
				AAT	69.4 %
Nucleotide 2649	G/A	Thr (348)	ACG/ACA	ACG	85.3 %
				ACA	14.7 %

The allele frequencies were based on analysis of 20 individuals.

EMBL ACCESSION NO: U 95626:

Nucleotide 40915	A/T Allele Frequency	A	97%
		T	3%

Nucleotide 41047 Insertion of ACA

ACAGCA	Allele Frequency	67%

ACA.(ACA) GCA		Allele Frequency		33%
Nucleotide 41058	C/A	Allele Frequency	C	82%
			A	18%
Nucleotide 41507	C/A	Allele Frequency	C	27%
			A	73%
Nucleotide 41768	A/T	Allele Frequency	A	7%
			T	93%
Nucleotide 42401	A/G	Allele Frequency	A	90%
			G	10%

Nucleotide 42598 Insertion of T

TTTCAA.....	Allele Frequency	85%
TTT(T)CAA...		15%

Nucleotide 42673	G/A	Allele Frequency	G	15%
			A	85%
Nucleotide 42723	C/A	Allele Frequency	C	70%
			A	30%

Nucleotide 42874	A/G	Allele Frequency	A	17%
			G	83%
Nucleotide 43018	A/T	Allele Frequency	A	87%
			T	13%

Allele frequencies were determined in a panel of 20 individuals

Example 2

Diagnostic assays for polymorphisms within the coding region of the CCR2 gene

The CCR2 gene has been cloned and published (EMBL Accession Number U80924 5471 bp) and all positions herein relate to the position therein unless stated otherwise or apparent from the context.

Methods: DNA preparation and Template preparation were performed as described above.

Nucleotide 2385 C/T Asn 260 AAC/AAT

Primary Fragment	Forward Oligo	Reverse Oligo	Annealing Temp	Time
1478-3544	1478-1497	3515-3534	55°	30s

Secondary Fragment	Forward Oligo	Reverse Oligo	Annealing Temp	Time
2111-2491	2111-2130	2472-2491	55°	30s

Polymorphic variation at position 2385 creates an SspI recognition sequence (AAT ATT), digestion of the PCR product with the restriction enzyme, SspI (New England Biolabs) can therefore distinguish the two polymorphic variants. Digestion of PCR products was performed according to the manufacturers instructions (New England Biolabs).

...TAT.AAC.ATT... wild type uncut

... TAT.AAT.ATT... variant cut

The size of full length, uncut PCR product is 380 bp (homozygous wild type), digestion of the PCR product from a homozygous variant will generate products of 275 bp and 105 bp, while digestion of the PCR product from a heterozygote will generate products of 380 bp, 275 bp and 105 bp.

Nucleotide 2649 G/A Thr (348) ACG/ACA

Methods (DNA and template preparation were performed as hereinbefore described)

	Primary Fragment	Forward Oligo	Reverse Oligo	Annealing Temp	Time
	1478-3544	1478-1497	3515-3534	55°	30s

Diagnostic Primer 2628-2648 TGGAGTGACTTCAACAAACAG (SEQ ID No: 1)

The diagnostic primer contains a single mismatch from the wild type sequence at the 3' residue (C->G).

Constant Primer 2821-2841 CATTGGGTGACATAGTCTGTA (SEQ ID No: 2)

PCR amplification using these primers will generate a product of 213 bp. The use of the diagnostic primer on a wild type template creates a StuI recognition sequence (AGGCCT) at the site of the polymorphism.

.....AGGCCT.... Wild type sequence

.....AGACCT.... Variant sequence

When a PCR product generated from a wild type template is digested with StuI (New England Biolabs), products of 191 bp and 22 bp will be generated. A product from a homozygous variant template will be uncut while digestion of a product from a heterozygous variant template will give rise to products of 215 bp, 195 bp and 20 bp.

Diagnostic Assays for polymorphisms in the promoter region of the CCR-2 gene

The human BAC (Bacterial Artificial Chromosome) clone 110P12 (Research Genetics) contains the CCR2 gene. The complete sequence of clone 110P12 has been published (EMBL Accession Number U95626,143068 bp) and all positions herein relate to the position therein unless stated otherwise or apparent from the context.

Nucleotide 42673

Variation at position 42673 (GCATG/A C) modifies an SphI recognition site (GCATGC). The PCR product (443 bp) containing the wild type sequence will be cleaved by SphI (New England Biolabs) to generate products of 189 bp and 254 bp. The homozygous variant product will be uncut (443 bp) while digestion of the heterozygote variant sequence will generate products of 189 bp, 254 bp and 443 bp.

Product	Forward Oligo	Reverse Oligo	Annealing temp	Time
42484-42927	42484-42503	42908-42927	58°	60s

Diagnostic Assays for polymorphisms in the promoter sequence of the CCR-2 gene employing Engineered Restriction Sites

Nucleotide 40915

Engineering of nucleotide 40917 (T-C) creates a Sau3A site (GATC). The wild type sequence

(GATC) is cleaved by Sau3A while the variant sequence (GTTC) is uncleaved.

Diagnostic Primer 40917-40937 5' CTGCAGTCTCAACCTCAAGCG (SEQ ID No: 3)

Constant Primer 40716-40736 5' CTGACTGAAATCTGGGCTGGG (SEQ ID No: 4)

Digestion of the PCR product (221 bp) containing the wild type sequence with Sau3A (New England Biolabs) will generate products of 24 bp and 197 bp. The variant sequence will be uncleaved while digestion of the heterozygous sequence will generate products of 221 bp, 197 bp and 24 bp.

Nucleotide 41507

Engineering of position 41505 creates an Acl I site (AACGTT), polymorphic variation at position 41507 (C/A) will modify this recognition sequence. The wild type sequence (AACGTT) will be cleaved while the variant sequence (AAAGTT) will not be cleaved.

Diagnostic Primer 41485-41505 5' CTGAGGTTCTTCTTGCTAAGA (SEQ ID No: 5)

Constant Primer 41695-41715 5' TAGGATTACAGGTGTGTGCCA (SEQ ID No: 6)

Digestion of the PCR product (230 bp) containing the wild type sequence with Acl I (New England Biolabs) will generate products of 208 bp and 22 bp. PCR products containing the homozygous variant sequence will be uncut while digestion of hererozygous products will generate products of 230 bp., 208 bp and 22 bp.

Nucleotide 41768

Engineering of position 41766 (G-C) creates an Nhe I recognition sequence (GCTAGC) and polymorphism (A/T) at position 41768 will modify this recognition sequence. The wild type

sequence (GCTAGC) will be cleaved while the variant sequence (GCTTGC) is not.

Diagnostic Primer 41746-41766 5' CTTGAACTCAGAAGGTGGAGC (SEQ ID No: 7)

Constant Primer 41968-41988 5'ACTTGGAGCAGAGCACCAGCA (SEQ ID No: 8)

Digestion of a PCR product (242 bp) containing the wild type sequence (GCTAGC) with Nhe I (New England Biolabs) will generate products of 22 bp and 220 bp. PCR products containing the variant sequence (GCTTGC) will be uncut while digestion of heterozygote products will generate products of 242 bp,220 bp and 22 bp.

Nucleotide 42401

Engineering of position 42403 (G-T) creates an Apo I recognition sequence (PuAATTPy). Polymorphic variation at position 42401 (A/G) will modify the recognition sequence. The wild type sequence (AAATTT) is cleaved while the variant sequence (AAGTTT) is not.

Diagnostic Primer 42403-42423 5' GTGGGTCTTTATACCTGGAAA (SEQ ID No: 9)

Constant Primer 42122-42142 5' TGCACAATGTATACATGTAGC (SEQ ID No: 10)

Digestion of a PCR product (301 bp) containing the wild type sequence (AAATTT) with Apo I will generate products of 23 bp and 278 bp. PCR products containing the variant sequence (AAGTTT) will be uncut while digestion of heterozygote products will generate products of 301 bp, 278 bp and 23 bp.

Nucleotide 42723

Engineering of position 42724 (A-G) will create a Mae II recognition sequence (ACGT). Polymorphic variation at position 42723 (C/A) will modify this recognition sequence. The wild type sequence (ACGT) is cleaved while the variant sequence is not.

Diagnostic Primer 42724-42744 5' TGTCTCAGGCAGTCCTGGTAC (SEQ ID No: 11)

Constant Primer 42541-42561 5' CCAATGTACAATGTTCCTGAC (SEQ ID No: 12)

Digestion of a PCR product (203 bp) containing the wild type sequence will generate products of 22 bp and 181 bp. Products containing the variant sequence are uncut while digestion of heterozygote products will generate products of 203 bp, 181 bp and 22 bp.

Nucleotide 43018

Engineering of nucleotide 43021 (C-A) creates a Bsr GI recognition sequence (TGTACA). The wild type sequence (AGTACA) is not cleaved while the variant sequence (TGTACA) is cleaved.

Diagnostic Primer 43019-43041 5' AGTGAGCCCCTGGGTGTGTGTAC (SEQ ID No: 13)

Constant Primer 42847-42866 5' GAACTGCAAAGCCTGCACAC (SEQ ID No: 14)

A PCR (197 bp) product containing the wild type sequence will not be cleaved by the restriction enzyme Bsr GI (New England Biolabs). Digestion of a PCR product containing the variant sequence will generate products of 172 bp and 25 bp, digestion of a heterozygote product will generate products of 197 bp, 172 bp and 25 bp.

Example 3

Haplotypes at positions 41507 and 42673

A double ARMS™ assay on the polymorphisms at 41507 (C/A) and 42673 (G/A) has generated haplotypes at these 2 positions as set out below.

41507 (C/A)	42673 (G/A)	
A	A	63 %
C	G	24 %

C.....A	13 %
A.....G	None Observed in 23 Individuals

Sequence Listing Free Text

Description of the artificial sequences in SEQ ID No: 1,3,5,7,9,11 and 13 is "PCR primer engineered RFLP".